Response to reviewers Round 3: Mutualists construct the ecological conditions that trigger the transition from parasitism

Ledru et al.

We would like to thank the reviewer for this third round of review which enabled us to improve the overall quality of our manuscript. We mainly modified the models in Appendix A2 and we removed some irrelevant sections. Below, you will find the reviewers' comments in black, and our responses in blue. The line numbers in our responses refer to the revised manuscript.

1 Reviewer, Eva Kisdi

At this time, I checked only the points I raised in the previous round.

Mathematical models: Appendix A2

My main concern is still Appendix A2. The authors have corrected equation (17) with the appropriate Poisson probability. However, this equation is still unclear to me at two points:

(i) It seems that mortality is the first to happen in the life cycle. If so, then a factor (1-m) should multiply ρ_{hs} in the bracketed expression (second line) of equation 17, because the host enjoys fecundity $f^h(\alpha_h, \alpha_s)$ only if its symbiont is present. The remaining fraction m of ρ_{hs} are hosts who have lost their symbionts and therefore should reproduce with fecundity f^{ha}.

(ii) The term $m(1-m)\rho_{hs}$ is the fraction of sites where the symbiont is lost. Because mortality occurs first, these hosts can be re-infected with the symbiont still before the next census; this is missing from both equations in (17). I am sorry for not having spotted this earlier.

The equation tacitly assumes that if both a host and a symbiont arrive in an empty cell, only the host survives. - It would be good to offer some help for the reader to interpret the terms of equation 17.

We address those issues by reformulating our model as follows (see the main text for a detailed description of each terms)

$$
\begin{aligned}
\rho_{ha}(t+1) \quad &= \Big\{ (1-m)\big(\rho_{ha}(t) + m\rho_{hs}(t)\big) + \big(1 - (1-m)\rho_h(t)\big)\Big(1 - \big((1-m)\rho_h(t)\big)^{\gamma_C}\Big) \\
&\quad \Big[1 - \exp\Big(-(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h, \alpha_s)\rho_{hs}(t)\big)\Big)\Big] \Big\} \\
&\quad \exp\Big(-(1-m)f^s(\alpha_h, \alpha_s)\rho_{hs}(t)\Big) \\
\rho_{hs}(t+1) \quad &= (1-m)^2\rho_{hs}(t) \\
&\quad + \Big\{ (1-m)\big(\rho_{ha}(t) + m\rho_{hs}(t)\big) + \big(1 - (1-m)\rho_h(t)\big)\Big(1 - \big((1-m)\rho_h(t)\big)^{\gamma_C}\Big) \\
&\quad \Big[1 - \exp\Big(-(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h, \alpha_s)\rho_{hs}(t)\big)\Big)\Big] \Big\} \\
&\quad \Big(1 - \exp\Big(-(1-m)^2 f^s(\alpha_h, \alpha_s)\rho_{hs}(t)\Big)\Big)
\end{aligned}
\tag{1}
$$

We do agree that after the mortality process the number of host alone should be $(1-m)\big(\rho_{ha}(t)+m\rho_{hs}(t)\big)$ to take into account: death of host alone and death of symbiont but survival of host with it. This discrepancy between hosts with alive symbiont and hosts with dead symbiont should also affect the reproduction. We modify the number of offspring that are present in the exponential terms.

In addition, we modify dispersal of symbiont. In our stochastic model, host disperse and settle and then symbiont do the same. So, when symbiont offspring disperse they can settle on alone host already

1

present in the population but also on new offspring. Thus the exponential term $\left(1 - \exp\left(-(1-m)^2 f^s(\alpha_h,\alpha_s)\rho_{hs}(t)\right)\right)$ is outside the bracket. And the host and offspring does not arrive in empty cell at the same time.

In order to explain in detail those mechanism, we add comments on the model in the Appendix A.2

line 704 page 22:

"In order to provide some heuristics about our stochastic model, we develop a simple deterministic model, with a monomorphic population of host and symbiont. The mathematical analysis also provides quantitative insights on our choice of parameters. More precisely, we aim to describe the expected proportion of sites occupied by a monomorphic population of hosts and symbionts at equilibrium with interaction traits α_h and α_s respectively. We assume no mutations of interaction or dispersal traits and hosts and symbionts disperse globally randomly over the landscape composed of N sites. According to our stochastic model (see Fig. ??), the dynamics of the proportion of sites occupied by the host alone ρ_{ha} or host with symbionts ρ_{hs} is given by

$$
\begin{aligned}
\rho_{ha}(t+1) \quad &= \Big\{ (1-m)\big(\rho_{ha}(t) + m\rho_{hs}(t)\big) + \big(1 - (1-m)\rho_h(t)\big)\Big(1 - \big((1-m)\rho_h(t)\big)^{\gamma_C}\Big) \\
&\quad \Big[1 - \exp\Big(-(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h,\alpha_s)\rho_{hs}(t)\big)\Big)\Big] \Big\} \\
&\quad \exp\Big(-(1-m)f^s(\alpha_h,\alpha_s)\rho_{hs}(t)\Big) \\
\rho_{hs}(t+1) \quad &= (1-m)^2\rho_{hs}(t) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2) \\
&\quad + \Big\{ (1-m)\big(\rho_{ha}(t) + m\rho_{hs}(t)\big) + \big(1 - (1-m)\rho_h(t)\big)\Big(1 - \big((1-m)\rho_h(t)\big)^{\gamma_C}\Big) \\
&\quad \Big[1 - \exp\Big(-(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h,\alpha_s)\rho_{hs}(t)\big)\Big)\Big] \Big\} \\
&\quad \Big(1 - \exp\Big(-(1-m)^2 f^s(\alpha_h,\alpha_s)\rho_{hs}(t)\Big)\Big)
\end{aligned}
$$

where $\rho_h = \rho_{ha} + \rho_{hs}$ is the total proportion of hosts. Since hosts and symbionts first face mortality with rate m, the proportion of host alone becomes $(1-m)\big(\rho_{ha}(t) + m\rho_{hs}(t)\big)$, where $m\rho_{hs}$ corresponds to hosts that have lost their symbiont, and the proportion of hosts with symbionts is $(1-m)^2\rho_{hs}(t)$. Then hosts produce offspring at a rate that depends on their partner: $f^{ha}(\alpha_h)$ (host alone) or $f^h(\alpha_h,\alpha_s)$ (host with symbiont). The total number of offspring is thus given by

$$
(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h,\alpha_s)\rho_{hs}(t)\big)
$$

Since offspring are dispersed randomly uniformly over the landscape, the probability that at least one offspring enters a cell is given by

$$
\Big[1 - \exp\Big(-(1-m)\big(f^{ha}(\alpha_h)(\rho_{ha}(t) + m\rho_{hs}(t)) + (1-m)f^h(\alpha_h,\alpha_s)\rho_{hs}(t)\big)\Big)\Big]
$$

However, they can only colonise empty cells, whose proportion is $(1 - (1-m)\rho_h)$. Moreover, once they enter an empty cell, their probability to establish in this cell depends on the host density ρ_h and it is given by $\Big(1 - \big((1-m)\rho_h(t)\big)^{\gamma_C}\Big)$. Finally, the symbionts produce offspring at a rate $f^s(\alpha_h,\alpha_s)$. Their offspring can only colonise alone host whose proportion is now given by the term between brackets. Since symbiont offspring are also randomly dispersed, the probability to invade a host alone is given by $\Big(1 - \exp\Big(-(1-m)^2 f^s(\alpha_h,\alpha_s)\rho_{hs}(t)\Big)\Big)$.

Equation (19) [formerly (20)] has also been revised, but unfortunately is not correct: When both types of symbionts arrive at a host, the probability that type m establishes is not the fraction of fecundity-weighted symbiont densities as it appears in equation 19 (or in equation 5 in the author's response). Instead, one has to calculate the probability that k of type m and j of type p arrives (product of two Poisson factors) and in this case, the probability of m establishing is $k/(k+j)$; then add up for all k and j. In other words, one needs to take the average of $k/(k+j)$ and not the quantity $(average\,k)/(average\,k + average\,j)$ as now. The two are not the same (Jensen's inequality).

It still seems to me that omitting the mathematical analysis would be the best option. Of course I like analytical results. However, the analysis needs to assume global dispersal of all hosts and symbionts

(even a pair approximation of the lattice model would be a taxing task). Since the main hypothesis is that local dispersal and clustering is at the heart of transition to mutualism, a model which cannot capture clustering is of limited interest.

We agree with the remarks on the Eq.(19) and (20). We did modify the model, however, the model becomes too complex and we were not able to provide any analytical insights. Moreover, as highlighted by the reviewer, we do not take into account explicit dispersal and explicit space, which is the key process in our stochastic model. Thus, we decide to remove further mathematical approximations of our stochastic model because they are irrelevant or inefficient to provide simple insights.

Other comments (numbering as in my previous review)

Annual growth rate and R_0

1. About R_0 of the host: $(1 - m)(1 + f^a)$ is the annual growth rate of the host, not R_0. R_0 is the lifetime number of offspring without competition, which is $(1 - m)f^a/m$. To calculate R_0, we start with a newborn at census. The first event is mortality, the newborn survives to first reproduction with probability $1 - m$. After this, it will reproduce on average $1 + (1 - m) + (1 - m)^2 + ... = 1/m$ times, each time producing f^a newborns. $R_0 > 1$ if and only if the annual growth rate > 1, but the two are not the same.

Sorry for the misuse of the notation R_0, we have modified the main text as follows to avoid any ambiguity:

line 206 page 5:

"... in the absence of mutation, the per capita growth rate at low densities of hosts is $F_0 = (1 - m)(1 + f^a)$. In our study, we have chosen parameters (see Table ??) for which F_0 is greater than 1 so that if we start with a large density of hosts initially, the probability of extinction is 0.

Selection strength

2. On γ_f being the selection strength on the trait α_s: the revision clarifies what the authors mean but "selection strength" is still inaccurate because γ_f relates to how the host fecundity increases with the symbiont trait α_s. The symbiont is not directly selected by the advantage of the host (although indirectly yes, but the indirect effect is not so simple as γ_f).

We agree that γ_f is not a direct selection strength but it indirectly influences the selection of symbionts. It is rather a mutualistic strength on host fecundity in the sense that it describes how fast the host fecundity increases with the symbiont trait. We thus modified the main text as follows

line 173 page 4:

"The parameter γ_f describes how fast the fecundity of the host increases with the interaction trait α_s of the symbiont. Since we are interested in the emergence of mutualism, the parameter γ_f describes the mutualistic strength of the symbiont on the host fecundity. In our model, we set the mutualistic strength of the host on the symbiont fecundity to $\gamma_s = 1$."

Competition strength

The following comments in my previous review (with line numbers referring to v2) went unnoticed: "In line 231, it is unclear why density is high when competition is strong. This is important to understand because the ms discusses the effect of competition strength but it is more likely the effect of density. In line 262, the % mutualistic symbionts decreases and not increases with γ_c. At this point, I was lost."

We previously answered this comment as follows: The parameter γ_C corresponds actually to the **inverse of the competition strength** for resources, we apologize for this mistake. That is why competition is strong when $\gamma_C < 1$ and weak when $\gamma_C > 1$. Indeed, the establishment probability is increasing with respect to γ_C, that is for a given density of host ρ, the establishment probability of an offspring goes to 0 when $\gamma_C \to 0$, while it increases to 1 when γ_C goes to ∞. Moreover, when $\gamma_C > 1$ the establishment probability decreases sub-linearly with the host density. Thus a slight increase in density reduces strongly the establishment probability. Inversely, when $\gamma_C > 1$, the relation is super-linear. Thus an increase in density induces a slight reduction of the establishment probability. In Fig.5(a), the % of mutualistic symbiont actually increases with the competition strength $1/\gamma_C$. We added some sentences to clarify this point and modify the Fig.5.

"The parameter γ_C corresponds to the inverse of the competition strength. Indeed, the establishment probability is increasing with respect to γ_C. Thus, the competition is strong when $\gamma_C < 1$ (sub-linear function), while it is weak when $\gamma_C \geq 1$ (super-linear function)."

Minor remarks: Recombination and transition

Speciation theory shows that recombination can prevent rather than only delay a transition like this. I do not understand why the authors are confident that recombination would only delay and not prevent the transition. (The last sentence of my previous review calling for the improvement of language was a general remark pertaining the whole paper, not only this sentence. Maybe the online system omitted a line break!)

In the related paragraph, we have added the following specification :

"In the present case a linkage disequilibrium between dispersal and interaction traits may also evolve".

We also did our best to improve language in the whole paper.